

Cell Chemical Biology

Tagging and Enriching Proteins Enables Cell-Specific Proteomics

Highlights

- A tetrazine-biotin probe containing a cleavable linker was created
- Proteomes labeled with cyclopropene amino acids were enriched and identified
- Proteome coverage is increased by targeting the amino acids to multiple codons
- Cell-specific proteomics was accomplished in the fly

Authors

Thomas S. Elliott, Ambra Bianco,
Fiona M. Townsley, Stephen D. Fried,
Jason W. Chin

Correspondence

chin@mrc-lmb.cam.ac.uk

In Brief

Elliott et al. synthesize a cleavable tetrazine-biotin probe enabling identification of proteins that are tagged and enriched via stochastic orthogonal recoding of translation with enrichment (SORT-E). Proteome tagging at multiple codons increases coverage, and the authors demonstrate that SORT-E at several codons enables cell-specific proteomics in the fly.



Tagging and Enriching Proteins Enables Cell-Specific Proteomics

Thomas S. Elliott,¹ Ambra Bianco,^{1,2} Fiona M. Townsley,¹ Stephen D. Fried,¹ and Jason W. Chin^{1,*}

¹Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK

²Present address: AstraZeneca, Darwin Building, Unit 310 Cambridge Science Park, Milton Road, Cambridge CB4 0FZ, UK

*Correspondence: chin@mrc-lmb.cam.ac.uk

<http://dx.doi.org/10.1016/j.chembiol.2016.05.018>

SUMMARY

Cell-specific proteomics in multicellular systems and whole animals is a promising approach to understand the differentiated functions of cells and tissues. Here, we extend our stochastic orthogonal recoding of translation (SORT) approach for the co-translational tagging of proteomes with a cyclopropane-containing amino acid in response to diverse codons in genetically targeted cells, and create a tetrazine-biotin probe containing a cleavable linker that offers a way to enrich and identify tagged proteins. We demonstrate that SORT with enrichment, SORT-E, efficiently recovers and enriches SORT tagged proteins and enables specific identification of enriched proteins via mass spectrometry, including low-abundance proteins. We show that tagging at distinct codons enriches overlapping, but distinct sets of proteins, suggesting that tagging at more than one codon enhances proteome coverage. Using SORT-E, we accomplish cell-specific proteomics in the fly. These results suggest that SORT-E will enable the definition of cell-specific proteomes in animals during development, disease progression, and learning and memory.

INTRODUCTION

Biomolecules, including proteins (Lang and Chin, 2014b; Ngo and Tirrell, 2011), DNA (Buskamp et al., 2014; Rieder and Luedtke, 2014), RNA (Asare-Okai et al., 2014), lipids (Yang et al., 2012), and sugars (Patterson et al., 2014; Prescher and Bertozzi, 2006), may be tagged with bio-orthogonal groups for many applications, including imaging, identification, and synthetic control. Global tagging of biomolecules may be achieved by feeding suitably tagged biosynthetic precursors to the natural biosynthetic machinery of the cell; an approach that requires the natural machinery to tolerate the introduction of the bio-orthogonal tag (Kiick et al., 2002). Analogs of natural precursors bearing bio-orthogonal tags may also be incorporated by engineered biosynthetic machineries with expanded substrate scope (Mahdavi et al., 2016; Ngo et al., 2013; Yuet et al., 2015). The creation of orthogonal biosynthetic machinery in cells enables the incorporation of diverse substrates without competition from endog-

enous substrates (Chin, 2014; Elliott et al., 2014a, 2014b; Liu and Schultz, 2010). In addition to their use for tagging biomolecules, bio-orthogonal groups have also found extensive utility in a variety of other approaches, including as reactive handles in activity-based probes of protein function (Cravatt et al., 2008).

The majority of approaches to bio-orthogonal labeling reported to date take advantage of copper (I)-catalyzed reactions between azides and alkynes (Hong et al., 2009; Wang et al., 2003). These components are relatively small, and so metabolic precursors modified with azides and alkynes may be tolerated by the natural biosynthetic machinery of cells. However, certain azides are prone to reduction in the cellular milieu (Mancuso et al., 2013), and the reaction between azides and linear alkynes depends on copper (I) catalysis, which requires optimization to avoid damage to biomolecules (Hong et al., 2009). Strain-promoted reactions between azides and strained alkynes remove the copper dependence (Agard et al., 2004). Although many of these reactions are relatively slow, varying the groups appended to the azide or strained alkyne has led to increases in reaction rates. However, despite the substantial impact of azide-alkyne cycloadditions on biological discovery, additional approaches for the labeling of biomolecules and probes are required.

Recently, inverse electron-demand Diels-Alder reactions between strained alkenes, alkynes, and tetrazines have been recognized as excellent bio-orthogonal reactions (Blackman et al., 2008; Devaraj et al., 2008; Lang et al., 2012b; Patterson et al., 2012; Yang et al., 2012). These reactions can be exceedingly rapid ($k \sim 10^6 \text{ M}^{-1} \text{ s}^{-1}$), proceed without catalysts, and produce nitrogen gas as the only by-product (Lang and Chin, 2014a). Alkenes and alkynes have been used to tag proteins, DNA, RNA, lipids, and sugars, enabling their labeling with tetrazine-fluorophore conjugates for imaging (Agarwal et al., 2015; Asare-Okai et al., 2014; Buskamp et al., 2014; Denk et al., 2014; Kurra et al., 2014; Lang et al., 2012a, 2012b; Nikic et al., 2014; Patterson et al., 2014; Pyka et al., 2014; Selvaraj et al., 2015; Uttamapinant et al., 2015; Yang et al., 2012). However, while azide-alkyne reactions form the basis of strategies to enrich tagged biomolecules for their identification by mass spectrometry (MS) (Bagert et al., 2016; Chesarino et al., 2014; Dieterich et al., 2006, 2007; Kleiner et al., 2015; Smeekens et al., 2015; Vanbeselaere et al., 2012; Zhang et al., 2013; Zheng et al., 2013), there are currently no methods for enriching and identifying biomolecules labeled with strained alkenes or alkynes through inverse electron-demand Diels-Alder reactions.

We, and others, have demonstrated that norbornene- (Kaya et al., 2012; Lang et al., 2012a; Plass et al., 2012), bicyclononyne- (Borrmann et al., 2012; Lang et al., 2012b),



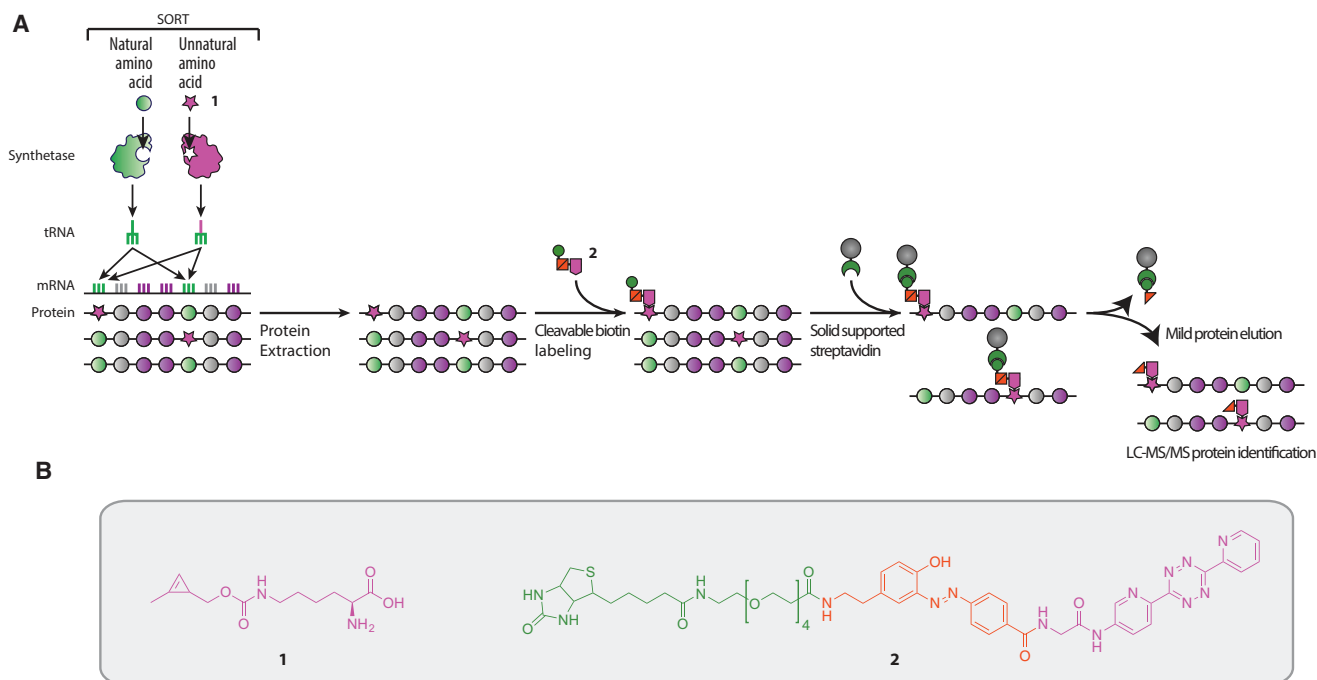


Figure 1. SORT-E Extends Proteome Tagging and Labeling to Protein Capture and Enrichment via Inverse Electron-Demand Diels-Alder Reactions with TDB, 2

(A) In SORT an unnatural amino acid (pink star) is recognized by an orthogonal pyrrolysyl-tRNA synthetase, and used to aminoacylate the cognate tRNA_{XXX} bearing a sense-decoding anticodon. This leads to the substoichiometric incorporation of the unnatural amino acid in response to the targeted sense codons. This approach has been implemented with the unnatural amino acid 1. In SORT-E the proteins are extracted from cells and SORT-tagged proteins are captured with the tetrazine diazobenzene biotin compound, 2, before capture on streptavidin beads. The beads are washed and enriched proteins are specifically eluted for detection by MS.

(B) Structures of the unnatural amino acid *N*-ε-(((2-methylcycloprop-2-en-1-yl)methoxy)carbonyl)-l-lysine, 1, and the cleavable TDB (2), used in SORT-E.

trans-cyclo-octene- (Lang et al., 2012b; Plass et al., 2012), and 1,3-disubstituted cyclopropene-containing (Elliott et al., 2014b) amino acids can be site-specifically incorporated into proteins in response to an amber codon introduced into a gene of interest using the pyrrolysyl-tRNA synthetase (PylRS)/tRNA_{CUA} pair and its active-site derivatives. This has enabled the imaging and control of protein function in vivo through the labeling of the tagged protein with appropriately functionalized tetrazine conjugates (Lang et al., 2012a, 2012b; Plass et al., 2012; Tsai et al., 2015; Uttamapinant et al., 2015).

An emerging application of bio-orthogonal labeling is for the imaging, and in some cases identification, of proteins expressed in particular cells at particular times in whole organisms (Elliott et al., 2014a, 2014b; Erdmann et al., 2015; Yuet et al., 2015). We reported the first solution to this problem: stochastic orthogonal recoding of translation (SORT; Figure 1A) in which the CUA anticodon of tRNA_{CUA} is converted to a variety of triplets (XXX) that are complementary to diverse sense codons. This approach enables the incorporation of diverse amino acids in response to diverse sense codons (Elliott et al., 2014b). We demonstrated that 1 is a substrate for PylRS, and can be used for SORT (Elliott et al., 2014b). Because SORT uses an orthogonal synthetase and tRNA, there is no competition for the active site of the synthetase between 1 and natural substrates. Thus, SORT can be used to label newly synthesized proteins in cells and animals without the use of minimal media or starvation. Moreover, because the

approach is genetically targeted it can be used to fluorescently label and identify newly synthesized proteins from specific cells, at specific developmental stages, within an animal via SORT with modification (SORT-M), in which 1 is labeled with tetrazine-fluorophore conjugates via an inverse electron-demand Diels-Alder reaction (Elliott et al., 2014b). We hypothesized that directing SORT to distinct codons by the use of tRNAs with distinct anticodons would lead to the labeling of different proteins with different efficiency, and that combining information on labeling at different codons may enable greater coverage of the proteome than labeling at any one codon (Elliott et al., 2014b).

Here we report SORT-E (SORT with enrichment) for the covalent capture and enrichment of SORT-labeled proteins via an inverse electron-demand Diels-Alder reaction with a tetrazine probe (Figure 1A). We demonstrate that SORT-E allows the substantial enrichment of proteomes tagged with 1 in response to diverse codons, and enables the identification and quantification of enriched proteins by MS (Bantscheff et al., 2012). We find that SORT-E does not preferentially identify proteins by molecular weight, but shows a slight bias toward the identification of low-abundance proteins, which should aid their identification. SORT-E at different codons leads to the enrichment of many proteins with different efficiencies, suggesting that proteome coverage may be increased by performing SORT experiments with several anticodon variants. To demonstrate the utility of SORT-E in a multicellular system, we create flies that tag newly

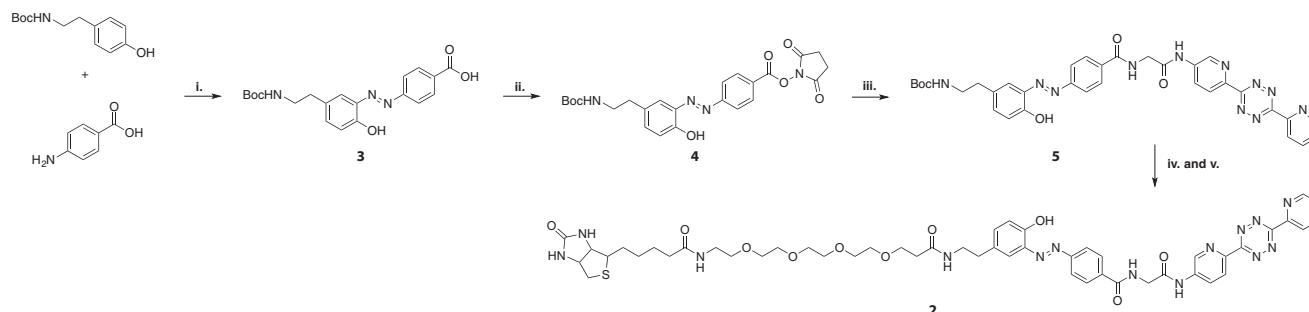


Figure 2. Synthesis of Compound 2

i. (a) 5 N HCl, NaNO₂, 0°C, (b) NaHCO₃, acetone, 0°C, 71% yield; ii. *N*-Hydroxysuccinimide, dimethylformamide (DMF), 1-ethyl-3-(3-dimethyl aminopropyl) carbodiimide, dimethylaminopyridine, room temperature, 48% yield; iii. Aminotetrazine, Et₃N, DMF, room temperature, 28% yield; iv. trifluoroacetic acid, CH₂Cl₂, room temperature, quantitative yield; v. NHS-PEG₄-Biotin (Thermo Scientific), DMF, Et₃N, room temperature, 18% yield.

synthesized proteins in ovary germ cells in response to several codons, and enrich and identify SORT-tagged proteins from these cells.

RESULTS

Design and Synthesis of a Cleavable Tetrazine Diazobenzene Biotin Probe

We designed a probe, tetrazine diazobenzene biotin (TDB, **2**; Figure 1B), to enable the enrichment, purification, and identification of molecules labeled with strained alkenes or alkynes that undergo inverse electron-demand Diels-Alder reactions with tetrazines. The design of our probe builds on previous work developing and optimizing reagents for the covalent capture of proteins through azide-alkyne cycloadditions (Yang et al., 2010). Our probe has three components: (1) a tetrazine, (2) a biotin moiety, and (3) a cleavable diazobenzene linker between the tetrazine and biotin moieties (Szychowski et al., 2010; Verhelst et al., 2007) (Figure 1B). The 3,6-dipyridyl tetrazine is designed to facilitate rapid and selective inverse electron-demand Diels-Alder reactions of the probe with the tagged molecules in a sample. This group has good stability, and we previously reported that a TAMRA derivative of this tetrazine reacts with **1** at position 150 in GFP with a rate constant of 27 M⁻¹ s⁻¹ (Elliott et al., 2014b). While azide-strained alkyne cycloaddition reactions have been reported that are only 10-fold slower than this reaction (Dommerholt et al., 2014), these reactions use azides coupled to aromatic rings with electron-deficient substituents that cannot be incorporated into proteins for proteome labeling. The rate constants reported for the reactions between cyclo-octynes and azides, which are similar to those used for proteome labeling (Dommerholt et al., 2014), are 100- to 1,000-fold lower than the rate constant for the reaction of **1** in GFP with 3,6-dipyridyl tetrazines. The biotin moiety in TDB was designed to enable the capture of the labeled proteins on streptavidin-coated beads and facilitate the removal of unlabeled molecules in the sample by washing. The diazobenzene was designed to enable the selective release of labeled molecules, via reductive cleavage, for identification by MS.

We developed a succinct synthesis of TDB probe **2** (Figure 2). Starting from two commercially available materials, a diazotization reaction afforded intermediate **3** in 71% yield. We found

that direct coupling of carboxylate **3** with previously synthesized tetrazine amines (Lang et al., 2012a) could not be achieved using conventional carbodiimide or HATU chemistry. However, reacting a tetrazine amine with an activated ester, **4**, gave the protected diazobenzene **5** in 28% yield. Boc deprotection of **5** followed by coupling with an activated biotin conjugate gave the desired product, **2**, in 18% yield over two steps.

Characterizing the Reaction of TDB with Proteins Incorporating **1**

Next we characterized (1) the labeling reaction between the probe and proteins in which we site-specifically incorporated **1**, and (2) the subsequent reductive cleavage of the linker by MS (Figure 3A). We produced T4 lysozyme (K-83-**1**)-His6, ubiquitin (K-6-**1**)-His6, and ubiquitin (K-48-**1**)-His6 from the corresponding genes (*T4 lysozyme* (83TAG)-His6, *Ubiquitin* (6TAG)-His6, and *Ubiquitin* (48TAG)-His6). Proteins were produced from *Escherichia coli* DH10B (T4 lysozyme) or BL21(DE3) (ubiquitin) expressing the PylRS/tRNA_{CUA} pair, which directs the incorporation of **1** in response to the amber codon, and purified by Ni-nitrilotriacetic acid chromatography with yields of 20–40 mg/l of culture. The incorporation of **1** in each protein was confirmed by electrospray ionization MS (ESI-MS) (Figures 3B–3D and S1).

We incubated each protein with 10 molar equivalents of **2** overnight at room temperature in 8 M guanidinium chloride. ESI-MS confirmed the quantitative labeling of each protein with **2** (Figures 3B–3D). Addition of Na₂S₂O₄ (25 mM, room temperature) to the labeled protein led to quantitative reductive cleavage of the diazo bond within 30 min, as judged by ESI-MS (Figures 3B–3D). We further analyzed the reduced products by liquid chromatography-tandem MS (LC-MS/MS) to directly confirm the identity and position of the final protein modification (Figure S2). Taken together, these experiments demonstrate that we can quantitatively and specifically capture biomolecules bearing 1,3-disubstituted cyclopropenes using TDB, and that we can quantitatively and specifically cleave the probe to an amino benzamide once it is ligated to the biomolecule of interest.

Enriching Labeled Proteins via SORT-E

Next we demonstrated that TDB (**2**) can be used to selectively enrich proteins that have been labeled with **1** via SORT. To label

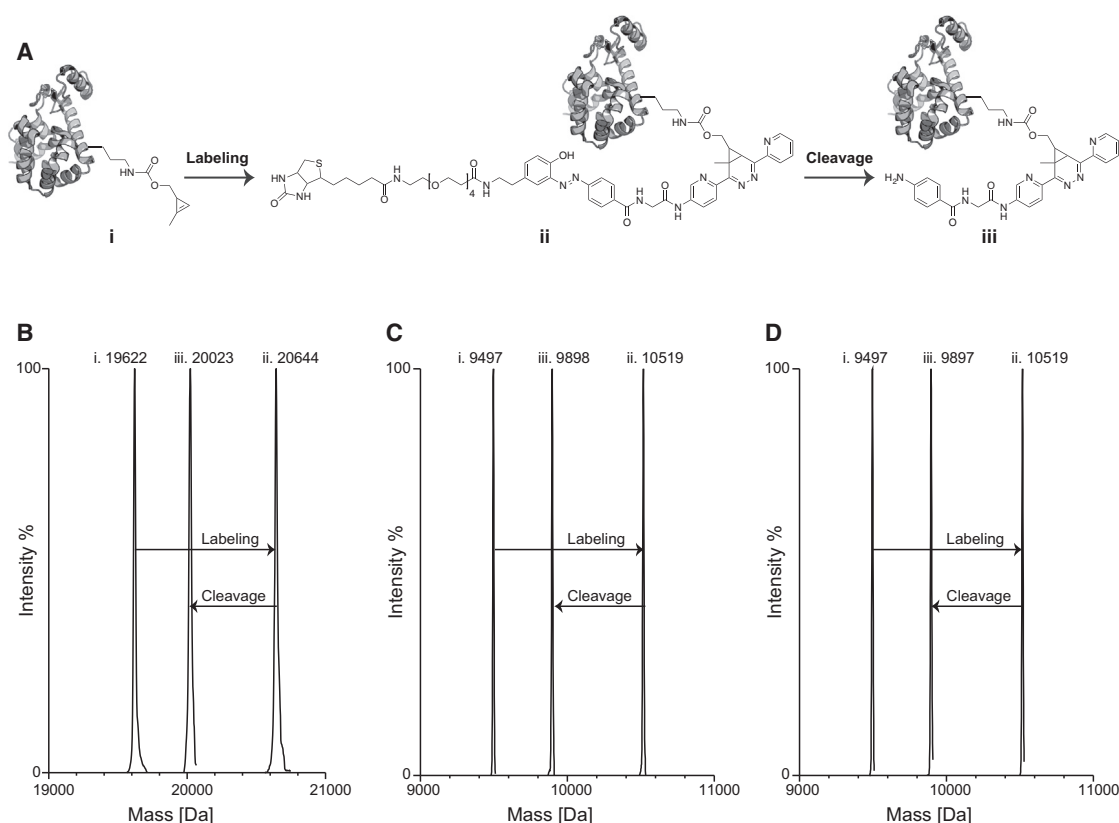


Figure 3. Quantitative Site-Specific Labeling of Genetically Encoded **1 with TDB, and Reductive Cleavage of the Diazobenzene by Mass Spectrometry**

(A) Proteins with site-specifically incorporated **1** (I) were labeled with 20 μ M **2** at room temperature overnight to furnish (II). Treatment of (II) with 25 mM $\text{Na}_2\text{S}_2\text{O}_4$ for 30 min generates the cleavage product (III).

(B) Deconvoluted mass spectra for species (I), (II), and (III) for T4-lysozyme (K-83-1)-His6.

(C) Deconvoluted mass spectra for species (I), (II), and (III) for ubiquitin (K-6-1)-His6.

(D) Deconvoluted mass spectra for species (I), (II), and (III) for ubiquitin (K-48-1)-His6. For each protein, mass increases by 1,022 Da upon conjugation with **2** (as expected), and then decreases by 621 Da upon reductive cleavage of the biotin moiety (as expected).

the proteomes of cells, we grew cultures containing one of four PylRS/tRNA_{XXX} pairs (where XXX = AGA, GCU, CAU, or UUU) in the presence of **1** (0.1 mM) (Elliott et al., 2014b). In control experiments we omitted **1** from the media. To confirm that **1** was incorporated into the proteome via SORT, we lysed the cells and treated a portion (240 μ g) of each lysate with a tetrazine-fluorophore conjugate **6** (Figure S3) (Lang et al., 2012a). As expected, we see fluorescent labeling of the proteomes derived from cells grown in the presence of **1**, but almost no labeling of the proteomes derived from cells grown in the absence of **1** (Figure S3).

To enrich SORT-labeled proteomes, we labeled protein lysates (500 μ l, at 8 mg/ml protein concentration) with **2** (20 μ M, final concentration) at room temperature overnight. The specific and rapid reaction of **2** with 1,3-disubstituted cyclopropenes allowed us to use a relatively low concentration of **2**. Using a low concentration of **2** allowed us to add the streptavidin-coated beads (that have a high capacity, 10 mg biotinylated protein per 1 ml of settled resin) directly to the reaction mixture, as the capture of conjugates between proteins and **2** by the streptavidin beads is not limited by the binding of free **2** to the streptavidin beads. This is in contrast

to other covalent enrichment approaches, which use much more labeling reagent in an effort to drive slower reactions to completion (equivalent azide reaction requires five times more linker, 1 mM CuSO_4 , 1 mM tris(2-carboxyethyl)phosphine, and an expensive ligand [Yang et al., 2010]) and precipitate proteins overnight to remove free label prior to capture on streptavidin. By omitting the precipitation step we simplify and accelerate the experiment, and by reducing the number of handling steps we potentially reduce protein loss and sources of experimental error. We incubated the labeled lysate with the beads (1.5 hr at room temperature) prior to washing.

We washed the beads extensively to remove non-specifically bound proteins. The final wash for all samples contained very little protein (Figure 4A), consistent with the effective removal of non-specifically bound proteins by the wash steps. Following the washes, we eluted proteins that were linked to the beads through the diazo group of **2** using PBS supplemented with 1% SDS and 25 mM $\text{Na}_2\text{S}_2\text{O}_4$. Substantial amounts of protein from samples grown in the presence of **1** were eluted with $\text{Na}_2\text{S}_2\text{O}_4$ while very little protein was eluted from samples grown in the absence of **1** (Figure 4A). We conclude that our procedure

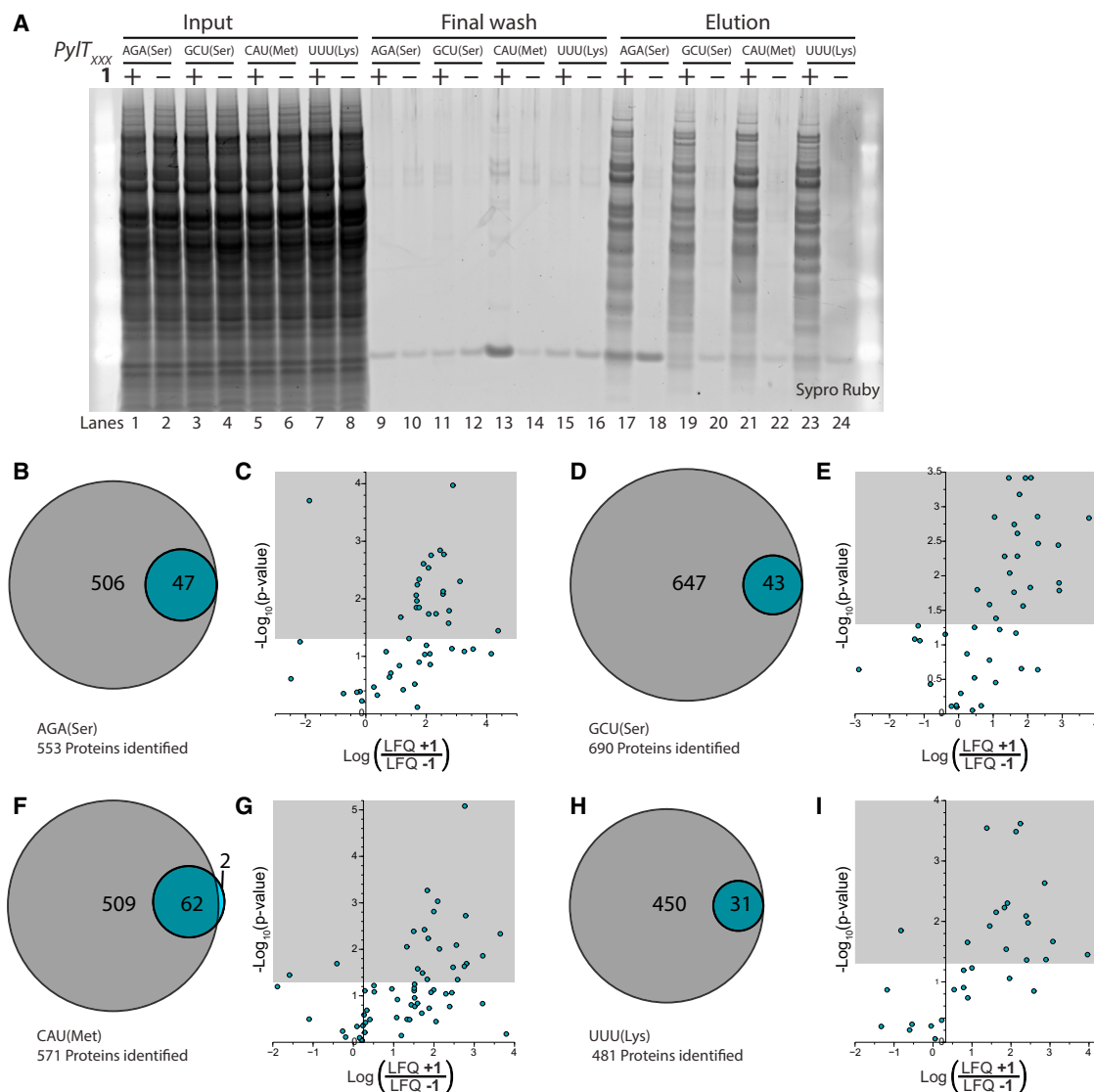


Figure 4. SORT-E Selectively Enriches Proteins Labeled with 1

(A) Cells expressing *PyIRS* and either *tRNA*_{AGA(Ser)}, *tRNA*_{GCU(Ser)}, *tRNA*_{CAU(Met)}, or *tRNA*_{UUU(Lys)} were grown either with or without **1** (0.1 mM). Lysates were labeled with **2** and tagged proteins captured with streptavidin beads, non-specifically bound proteins were washed away, and specifically captured proteins were eluted by cleavage of **2**. Aliquots from the initial lysate, the final wash, and eluted fraction were analyzed by SDS-PAGE.

(B, D, F, and H) Venn diagrams representing the number of proteins identified by MS from the SORT-E elution (gray) and the no-amino-acid control (blue).

(C, E, G, and I) Proteins that are identified in both the SORT-E samples (+1) and controls (-1) are significantly enriched in the +1 samples. The volcano plots show the log ratio of the LFQ values for each protein in this subset (i.e., the enrichment factor), plotted against the log p value of the null hypothesis that there is no difference between the LFQ values. The area shaded in gray corresponds to the threshold of p values of <0.05.

enables the specific capture of proteins labeled with **1**. Quantifying the protein in the GCU (Ser) elution (Figure S4) reveals that approximately 39 μ g of protein was recovered. Since the input for the pulldown was 4 mg of protein, and SORT tags targeted proteins in response to target codons with an efficiency of substantially less than 1% (Elliott et al., 2014b), this enrichment step efficiently captures tagged proteins. We conclude that our approach, which we call stochastic orthogonal recoding of translation with enrichment (SORT-E), efficiently recovers and enriches SORT-labeled proteins. To define the limits of specific enrichment, we mixed SORT-labeled lysates with increasing

amounts of unlabeled lysates. With a 10-fold excess of unlabeled lysate SORT-E still leads to selective enrichment, while for a 100-fold excess of unlabeled lysate non-specific labeling approaches the level of specific labeling (Figure S5). These data demonstrate that SORT-E provides a powerful approach for enriching proteins that are substoichiometrically tagged in target cells from a vast excess of untagged proteins.

Identifying Enriched Proteins by Mass Spectrometry

To demonstrate that our approach enables the specific identification of proteins labeled with **1** by SORT and to further

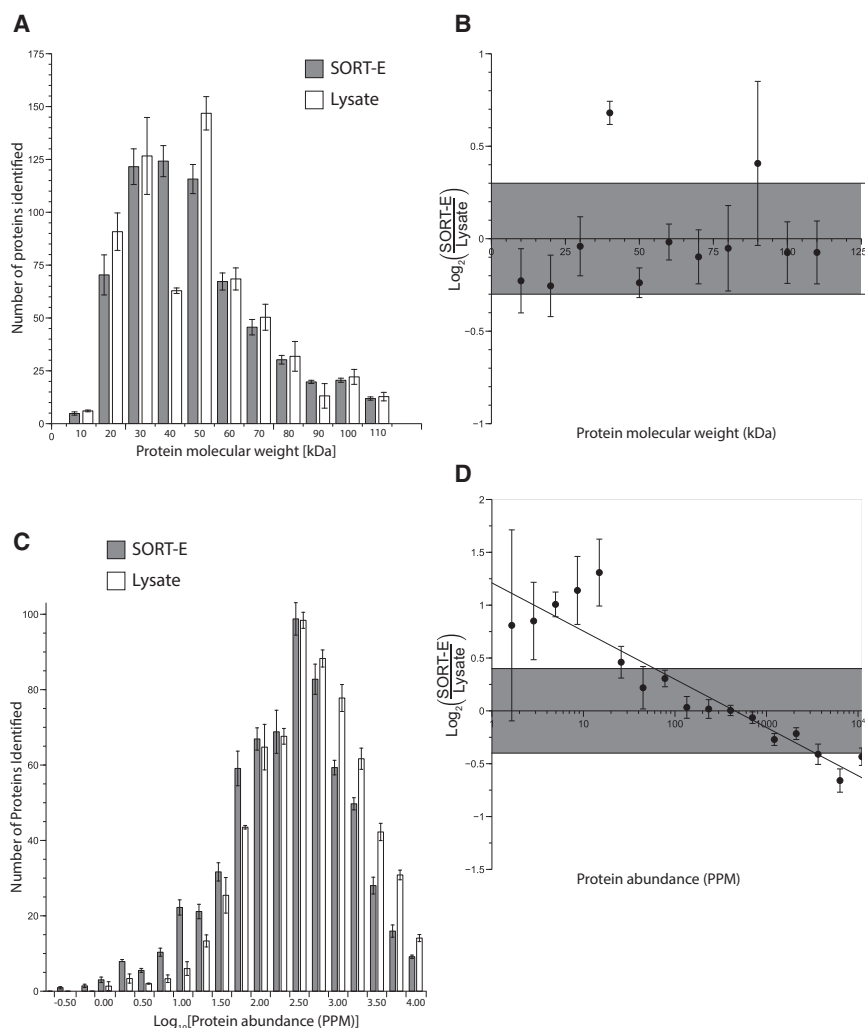


Figure 5. SORT-E Enriches Proteins of Most Molecular Weights Equally, but Preferentially Enriches Less Abundant Proteins from the Proteome

(A) Number of proteins identified, either from crude lysates (white bars) or from SORT-E (gray bars). Data are binned by molecular weight. Error bars represent the SD of three biological replicates.

(B) SORT-E identifies proteins across the full range of molecular weights of the *E. coli* proteome and shows no bias for larger or smaller proteins. The graph shows the log ratio of the number of proteins identified from SORT-E to the number of proteins identified from crude lysates, as a function of protein molecular weight, using the bins from (A). Error bars represent the SD of three biological replicates. Points that fall within the shaded area have a relative abundance of proteins before and after SORT-E that is within a factor of 1.32.

(C) Number of proteins identified, either from crude lysates (white bars) or from SORT-E (gray bars). Data are binned by protein abundance in the PAX database of the *E. coli* proteome. Error bars represent the SD of three biological replicates.

(D) SORT-E preferentially enriches low-abundance proteins. Log ratio of the number of proteins identified from SORT-E to the number of proteins identified from crude lysates, as a function of protein abundance defined in the PAX database. Error bars represent the SD of three biological replicates. Points that fall within the shaded area have a relative abundance of proteins before and after SORT-E that is within a factor of 1.32.

quantify the specificity of the approach for identifying labeled proteins, we performed in-gel tryptic digests and LC-MS/MS of the eluted proteins. For cells grown in the presence of **1** and the PylRS/tRNA_{XXX} pairs we identified 553, 690, 571, and 481 proteins (for XXX = AGA, GCU, CAU, and UUU, respectively) from approximately 1.3 μ g of protein (i.e., using just 3.3% of our elution) (Figures 4B–4I and Table S1). The samples from cells grown in the absence of **1** yielded a small subset of the proteins identified from cells grown in the presence of **1**, with two exceptions for CAU. We identified 47, 43, 64, and 31 proteins from cells grown in the absence of **1** (for XXX = AGA, GCU, CAU, and UUU, respectively, Figures 4B–4I). Thus 92% ($\pm 2.2\%$) of proteins identified from cells grown in the presence of **1** are not found in cells grown in the absence of **1**. For the remaining 8% ($\pm 2.2\%$) of proteins that are captured for cells grown in both the presence and absence of **1**, we compared the abundance of each protein, as judged by label free quantification (LFQ) (Figures 4C, 4E, 4G, 4I, and S6) (Cox et al., 2014). The vast majority of proteins that are found in both samples, and for which we have statistically significant data, are enriched in the cells grown in the presence of **1** ($\log[\text{LFQ} + 1/\text{LFQ} - 1] > 0$).

SORT-E Enhances the Identification of Low-Abundance Proteins

To address the extent to which our approach reproduces the distribution of proteins present in the proteome, we compared enriched proteomes identified via SORT-E with the proteome prior to TDB enrichment (Figure 5A). We find that the relative abundance of proteins for the vast majority of molecular weight ranges is comparable before and after TDB enrichment (Figures 5A and 5B). However, for the molecular weight range between 30 and 40 kDa we observed an increase in the number of proteins identified after the TDB enrichment.

To address the extent to which TDB enrichment captures protein present at different abundances in the cell, we compared the distribution of proteins before and after SORT-E as a function of abundance in the cell, using the data compiled in the PAX database (Figures 5C and 5D) (Wang et al., 2012, 2015). We found that the distribution of proteins following TDB enrichment (Figure 5C) is slightly shifted to favor low-abundance proteins, with respect to the distribution of proteins prior to enrichment. This observation was further supported by additional experiments (Figure S7). This basis of this enrichment is unclear, but it may result from the TDB enrichment decreasing sample complexity

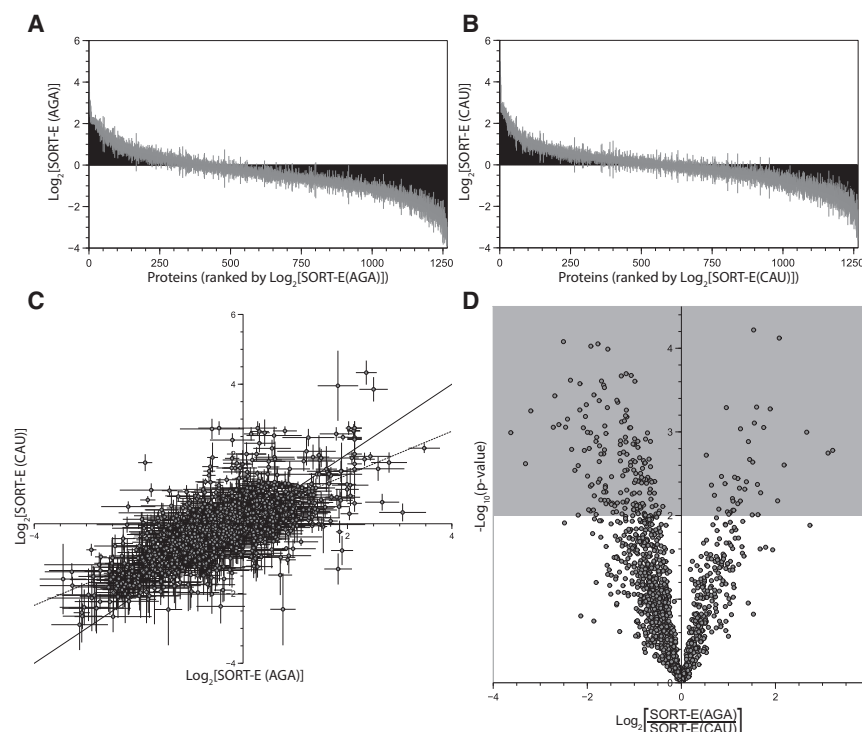


Figure 6. The Efficiency of Enrichment Depends on the Identity of the Anticodon used in SORT-E

(A) SORT-E (AGA, Ser) enriches different proteins in the proteome to different extents. The graph shows the log of the ratio of the relative abundance for proteins detected for SORT-E to the relative abundance prior to enrichment. The data were measured in a multiplexed TMT experiment in which 1,265 proteins were common to all samples. The data are ranked by this ratio from most enriched to least enriched, to show the distribution of enrichments. Error bars, shown in gray, are the SD of three biological replicates each performed in duplicate.

(B) As for (A), but for SORT-E (CAU, Met).

(C) SORT-E (AGA, Ser) and SORT-E (CAU, Met) enrichments are globally correlated. The values in (A) and (B) were plotted to examine the correlation between protein enrichments in SORT experiments that use tRNAs with two distinct anticodons (AGA and CAU). The correlation (least-squares regression line $y = 0.157 + 0.624x$; $R^2 = 0.48$, $p = 10^{-179}$) between the enrichments for each codon is indicated by the dashed line. The solid line represents $y = x$, indicating what the best fit would look like if the enrichments were identical between the two anticodons. Error bars represent the st. dev. of three biological replicates.

(D) SORT-E with distinct anticodons enriches a substantial subset of proteins with different efficiencies. The volcano plot shows the ratio of the

data shown in (C), plotted against the p value of the null hypothesis that the protein is enriched with identical efficiencies for SORT-E (AGA, Ser) and SORT-E (CAU, Met). The 203 points in the upper shaded region correspond to the 16% of proteins that are significantly ($p < 10^{-2}$) more enriched by SORT-E (AGA, Ser) (right branch) or SORT-E (CAU, Met) (left branch).

and bringing some low abundance proteins into a concentration range where they can be more easily detected by MS.

SORT-E at Different Codons Enriches Different Proteins

To investigate how the relative abundance of individual proteins is affected by the TDB pulldown using SORT systems addressed to different codons, we performed multiplexed tandem mass tagging (TMT) labeling and MS (McAlister et al., 2012; Thompson et al., 2003; Werner et al., 2012) (Figures 6 and S8; Table S2). For SORT-E (CAU, Met), in which 1 is stochastically incorporated in place of Met using a tRNA with a CAU anticodon, we observed proteins that increase as a fraction of total proteins after TDB enrichment, and proteins that are unchanged or depleted as a fraction of total proteins after TDB enrichment (Figure 6A). For SORT-E (AGA, Ser), in which 1 is stochastically incorporated in place of Ser using a tRNA with an AGA anticodon, we observed proteins that increase as a fraction of total proteins after TDB enrichment, and proteins that are unchanged or depleted as a fraction of total proteins after TDB enrichment (Figure 6B). We find that, broadly, there is a positive correlation in enrichment of proteins in SORT-E (CAU, Met) and SORT-E (AGA, Ser) (Figure 6C). However, there are clearly many proteins that are significantly and preferentially enriched by SORT-E (CAU, Met) over SORT-E (AGA, Ser) and vice versa (Figure 6D). The ratio of enrichments for SORT-E with CAU and AGA anticodons does not correlate with the relative abundance of their cognate codons in genes, as might be predicted by a simple model in which proteins are enriched on the basis of the relative frequency of a

target codon in a gene (Figure S9). Similarly, the ratio of enrichments does not correlate with the solvent accessibility of the amino acids targeted for replacement by SORT-E, as might be expected if solvent accessibility controlled the extent to which 1 is tolerated and/or labeled at site in a protein (Figure S10). It remains possible that a combination of these factors may determine enrichment efficiency.

SORT-E in a Multicellular System

We have previously demonstrated a modular system for SORT labeling in the fly *Drosophila melanogaster* (Elliott et al., 2014b). In this system *PyIRS* is expressed from a promoter that contains a GAL4 upstream activating sequence (*UAS-PyIRS*), and its cognate tRNA_{xxx} is expressed from a *PyT_{xxx}* gene on a U6 promoter. By crossing *UAS-PyIRS/PyT_{xxx}* flies with flies that express the GAL4 transcription factor in specific cells at specific times, we were able to effect cell-specific proteome labeling. For example, we previously demonstrated selective SORT-M labeling in the germ cells of the fly ovary from stage 5 of oogenesis onward, using *nos-vp16-GAL4, PyIRS, PyT_{UGC}* flies (Elliott et al., 2014b).

To demonstrate that we can extend SORT-E to multicellular systems, we performed SORT tagging and enrichment of the proteome from the germ cells of fly ovaries. To enhance proteome tagging efficiency, and potentially proteome coverage, we first created flies containing additional copies of *PyIT* with distinct anticodons. We found that the proteomes of flies containing two *PyIT* genes with distinct anticodons, drawn from the group *PyIT_{UGC}* (Ala) and *PyIT_{CAU}* (Met), *PyIT_{GCU}* (Ser), were

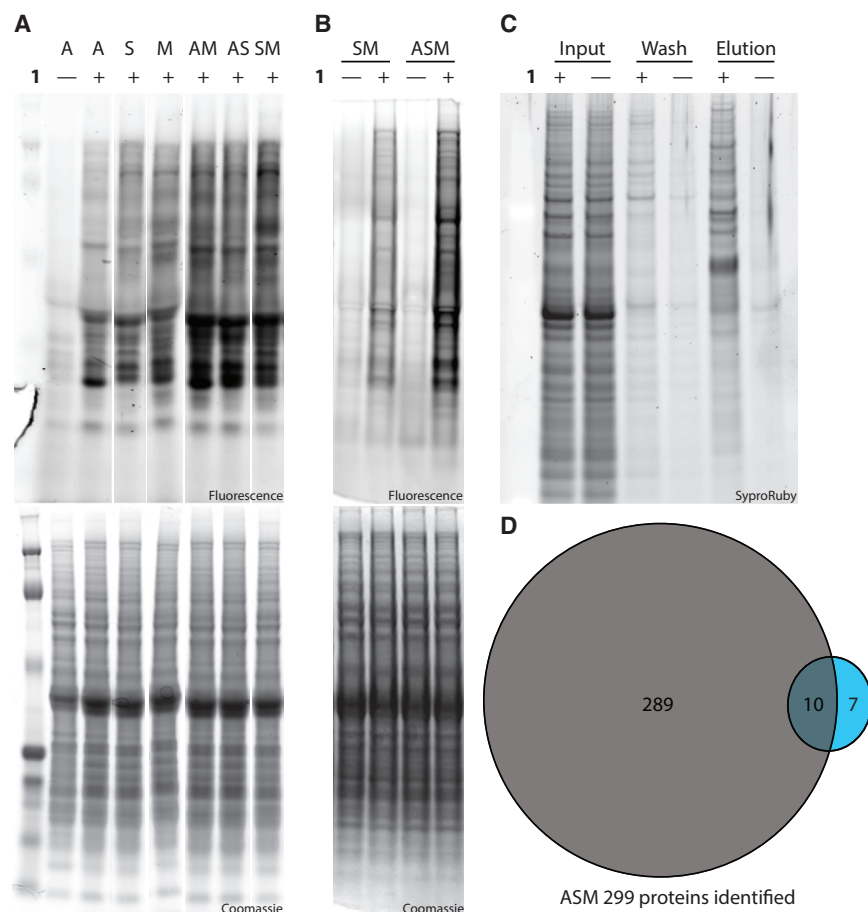


Figure 7. SORT-E from the Germ Cells of Fly Ovaries

(A) Proteome labeling from fly ovaries is enhanced by combining two *PyIT_{XXX}* genes with distinct anticodon sequences. Ovaries were harvested from 15 females expressing *PyIRS* and *PyIT_{XXX}* fed normal food containing **1** (10 mM), or in the control lane from 15 females not expressing *PyIRS* and *PyIT_{XXX}* but fed normal food containing **1** (10 mM). Protein lysates were labeled with tetrazine-fluorophore conjugate **7** (4 μ M, Figure S11). *PyIT_{XXX}* corresponds to variants of *PyIT* bearing anticodon A = *PyIT_{UGC}* (Ala); S = *PyIT_{GCU}* (Ser), and M = *PyIT_{CAU}* (Met) or combinations bearing two variants of *PyIT_{XXX}*, for example, AM = *PyIT_{UGC}* (Ala) and *PyIT_{CAU}* (Met). The data shown are from a gel in which intervening lanes have been removed (indicated by white space). The full gel is detailed in Figure S11.

(B) Proteome labeling from fly ovaries is further enhanced by combining three *PyIT_{XXX}* genes with distinct anticodon sequences. Experiment performed as in (A). The data shown are from a gel in which intervening lanes have been removed (indicated by white space). The full gel is detailed in Figure S11.

(C) SORT-E from the germ cells of *nos-vp16-GAL4*, *PyIRS*, *PyIT* (Ala, Met, Ser) fly ovaries. Ovaries from flies fed amino acid **1** or control flies not fed the amino acid were lysed. Lysates were labeled with **2** and tagged proteins captured with streptavidin beads, non-specifically bound proteins were washed away, and specifically captured proteins were eluted by cleavage of **2**. All steps were performed as described for *E. coli* without further optimization. Aliquots from the initial lysate, the final wash, and eluted fractions were analyzed by SDS-PAGE.

(D) Venn diagrams representing the number of proteins identified by MS from the SORT-E elution from flies fed **1** (gray) and the no-amino-acid control (blue).

more efficiently labeled than flies containing any one of these *PyIT_{XXX}* sequences alone (Figures 7A and S11). Moreover, we found that the proteomes of flies containing *PyIT_{UGC}* (Ala), *PyIT_{CAU}* (Met), *PyIT_{GCU}* (Ser) were more efficiently labeled than the proteomes of flies containing the most efficient pair of *PyIT_{XXX}* variants (Figures 7B and S11) and used these flies, henceforth referred to as *PyIRS*, *PyIT* (Ala, Met, Ser) flies, for further experiments.

nos-vp16-GAL4, *PyIRS*, *PyIT* (Ala, Met, Ser) flies were fed with or without the addition of **1** to their food to provide SORT-labeled flies and genetically matched control flies that were not labeled via SORT. The proteins from the ovaries of each group of flies were extracted and subjected to the TDB-mediated enrichment procedure described herein for *E. coli*, without further optimization. Using this approach we were able to selectively enrich the proteome from the germ cells of the ovaries of flies fed amino acid **1** (Figure 7C). Additional experiments suggested that using 2-fold more *nos-vp16-GAL4*, *PyIRS*, *PyIT* (Ala, Met, Ser) ovary lysate than *PyIRS/PyIT_{GCU}* *E. coli* lysate as an input yielded comparable amounts of specifically eluted proteins (Figure S12). This suggests that the efficiency of SORT tagging in this fly line

approaches that in *E. coli*. To further quantify the specificity of the approach for identifying labeled proteins from the fly ovary, we performed small-scale in-gel tryptic digests and LC-MS/MS on 2 μ g of the eluted proteins. We identified 299 proteins in flies fed amino acid **1** and 17 proteins in control flies that were not fed the amino acid (Figure 7D and Table S3). The proteins shared between the two samples, on which we have statistically significant data, show amino acid-dependent enrichment (Figure S13). These data clearly demonstrate that SORT-E can be used to enrich tagged proteins from a multicellular system.

DISCUSSION

We have developed a TDB probe for the covalent capture of proteins containing **1**.

We have taken advantage of the site-specific incorporation of **1**, via genetic code expansion, to carefully characterize both the reaction of this probe with proteins containing **1** at specific sites and the reductive cleavage of the probe, ensuring that both steps are quantitative. We have demonstrated that proteomes labeled with **1** via SORT at diverse codons can be

efficiently enriched (SORT-E) and then identified, demonstrating that inverse electron-demand Diels-Alder reactions can be used for the enrichment and identification, in addition to the imaging and control, of biomolecules. We anticipate that our approach may be extended to the capture, enrichment, and identification of proteins and other biomolecules that may be labeled with strained, or otherwise activated, alkenes or alkynes.

We demonstrate that SORT-E does not preferentially identify proteins of most molecular weights, and that SORT-E shows a slight bias toward the identification of low-abundance proteins, which should aid their identification. We find that SORT-E at different codons leads to the enrichment of many proteins with different efficiencies. While it has previously been argued that labeling at methionine is sufficient to cover the proteome (Yuet and Tirrell, 2014), as all proteins contain an N-terminal methionine we find that numerous proteins are more effectively enriched and detected via labeling at a non-methionine codon. Our results also suggest that coverage may be increased by tagging the proteome using tRNAs that target several codons. Moreover, we demonstrate that SORT-E can be used to selectively enrich the proteomes, tagged in response to several codons, from specific cells in the fly.

Because azide/alkyne cycloadditions and cyclopropene/tetrazine cycloadditions are mutually orthogonal bio-orthogonal reactions (Sachdeva et al., 2014; Shih et al., 2014), it will be possible to label distinct biomolecules in a single sample with azides or alkynes and cyclopropenes, and selectively and independently enrich and identify each population of labeled biomolecules. Given the diversity of biomolecules that can now be labeled with azides and strained alkenes, this opens up many exciting new experimental possibilities.

We anticipate that the combination of SORT, which we have previously demonstrated can be used to label the proteomes of selected cells within a whole organism, and the enrichment strategies we have carefully characterized here will enable the identification of proteins expressed in specific cells at specific times in development, disease progression, and learning and memory. Indeed, we are currently using SORT-E, in an expanding set of multicellular systems, to define cell-specific proteomes.

SIGNIFICANCE

Well-characterized methods for tagging and enriching the proteomes of genetically targeted cells will provide a foundation for cell-specific proteomics in multicellular systems and whole animals. Such approaches are essential for defining how the distinct proteomes of cells in a body, or other multicellular system, carry out differentiated functions. Co-translational labeling of proteomes in *E. coli* and genetically targeted cells in the fly with a cyclopropene group, coupled with covalent capture via a tetrazine-biotin probe containing a cleavable linker, enables us to enrich labeled proteins from genetically targeted cells in an approach we term stochastic orthogonal recoding of translation with enrichment (SORT-E). Our approach allows the tagging of proteins in response to distinct codons. We demonstrate that tagging at distinct codons leads to different enrichments, suggesting that labeling at more

than one codon enhances proteome coverage. We anticipate that SORT-E will enable the definition of cell-specific proteomes in animals during development, disease progression, and learning and memory.

EXPERIMENTAL PROCEDURES

Sources of chemicals, details of chemical synthesis and MS, data analysis, and computer code used for analysis are provided in [Supplemental Experimental Procedures](#).

SORT-E in *E. coli*

SORT-E involves chemoselective labeling of proteomes tagged with **1** by TDB conjugate **2**, and enrichment and elution of SORT-E proteins.

For capture, enrichment, and elution of tagged proteins, 500 μ l of cleared cell lysate (3.5–4 mg of 8 mg ml⁻¹ lysate) was typically used. To 500 μ l of cleared cell lysate, DTT (1 mM) was added and mixed with gentle vortexing, and the mixture was incubated at room temperature for 45 min. Iodoacetamide (5.5 mM) was added, mixed with gentle vortexing, then incubated for 30 min. **2** (5 μ l, 20 μ M final concentration, from a 2 mM stock in DMSO) was added, the reactions mixed by gentle vortexing, and the samples incubated in the dark overnight with end-over-end rotation (Hula mixer). 10- μ l aliquots were taken at this stage as input sample for subsequent SDS-PAGE analysis. The samples were then diluted with PBS to a final volume of 5 ml and high-capacity streptavidin beads (150 μ l of settled resin, pre-equilibrated in PBS, Thermo Scientific Streptavidin Agarose resin) added. The beads were incubated with end-over-end rotation (Hula mixer) for 1.5 hr and then collected by gravity filtration through a column (Poly-Prep chromatography column, Bio-Rad 731-1550). The beads were resuspended in urea buffer (500 μ l, 8 M urea, 25 mM Tris, [pH 8]) and transferred to a smaller spin column (Mini Bio-spin chromatography column, Bio-Rad 731-1550). The beads were collected by mild centrifugation (1,000 rpm, 5-s pulse) and washed 2 \times urea buffer (500 μ l) then 2 \times 1% SDS in PBS (500 μ l) with 15-min incubation times between each wash. The final wash acts as a control for the Na₂S₂O₄ specific elution; the beads were resuspended in 1% SDS in PBS (150 μ l) and incubated at room temperature for 30 min with end-over-end rotation (Hula mixer). The beads collected by mild centrifugation (1,000 rpm, 5-s pulse) and the supernatant kept for subsequent analysis by SDS-PAGE. Specifically bound proteins were then eluted by resuspending the beads in 1% SDS in PBS supplemented with 25 mM Na₂S₂O₄ (150 μ l) and incubating at room temperature for 30 min with end-over-end rotation (Hula mixer). The supernatant was collected by mild centrifugation (1,000 rpm, 5-s pulse) and analyzed by SDS-PAGE.

Fly Lines and Culture Conditions

All flies were grown at 25°C on standard Iberian medium. Flies were fed **1** by mixing dried yeast with a solution of **1** (10 mM) to form a paste. This paste was added as a supplement to the normal Iberian fly food for a minimum of 24 hr and the yeast was changed daily.

Double- and triple-sense codon lines were created by recombination using the original lines FT58 (A = *Pyl*_{T_{UGC}}, Ala), FT60 (S = *Pyl*_{T_{GCU}}, Ser), FT62 (L = *Pyl*_{T_{CAG}}, Leu), and FT63 (M = *Pyl*_{T_{CAU}}, Met) (Elliott et al., 2014b). Trans-heterozygous virgins were collected for each pairwise combination of sense codon (AS, FT58/FT60; AL, FT58/FT62; AM, FT58/FT63; SL, FT60/FT62; SM, FT60/FT63; and LM, FT62/FT63) and crossed to males of the third chromosome balancer stock w⁺;TM3/TM6. Potential recombinant males were identified based on eye color, and individuals were backcrossed to virgins of w⁺;TM3/TM6 to make a balanced stock. Recombinant lines were then screened by crossing to *nos-vp16-GAL4* virgins (Bloomington 4937) to create FT58-60/*nos-vp16-GAL4* (AS), FT58-62/*nos-vp16-GAL4* (AL), etc., and compared with the original single-sense codon lines FT58/*nos-vp16-GAL4*, FT60/*nos-vp16-GAL4*, FT62/*nos-vp16-GAL4*, and FT63/*nos-vp16-GAL4*. The females were fed 10 mM **1** for 24–48 hr, after which the ovaries were extracted from 15 females of the indicated genotype and labeled with 4 μ M **7** for 2 hr as described by Elliott et al. (2014b).

Triple-sense codon lines were generated in a similar manner. In this case the double-sense codon lines were crossed to a different single-sense codon line to generate trans-heterozygotes with three different sense codons. In this case

not all combinations produced viable trans-heterozygotes and in some cases trans-heterozygotes were viable but gave no potential recombinant males. Successful combinations were FT58-63 (AM)/FT60 (S), FT60-63 (SM)/FT62 (L), and FT58-60 (AS)/FT63 (M), which generated the lines FT58-60-63 (ASM) and FT60-62-63 (SLM). These lines were screened by crossing to nos-*vp16*-GAL4 virgins and compared with the double-sense codon lines as described above (Elliott et al., 2014b).

SORT-E from *D. melanogaster*

Ovaries were dissected from 250 female flies of FT58-60-63/*nos-*vp16*-GAL4* (ASM), which had been fed normal food either supplemented with **1** (10 mM) or without **1**. The ovaries were homogenized into 8 M urea and 15 mM Tris (250 μ l) and the resultant protein lysate clarified by filtration. A Bradford assay was used to determine the protein concentration. For SORT-E and subsequent MS, typically 7 mg of fly ovary protein was used and labeled with **2** (20 μ M) in an identical procedure as described in the protocol for *E. coli* above.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, 13 figures, three tables, and Supplemental Program and can be found with this article online at <http://dx.doi.org/10.1016/j.chembiol.2016.05.018>.

AUTHOR CONTRIBUTIONS

T.S.E. performed all chemical and biochemical experiments with *E. coli*. A.B., F.M.T., and T.S.E. performed the biochemical experiments with *D. melanogaster*. T.S.E. and S.D.F. analyzed the data. J.W.C. and T.S.E. wrote the paper with input from all authors.

ACKNOWLEDGMENTS

This work was supported by the Medical Research Council, UK (MC_U105181009 and MC_UP_A024_1008, to J.W.C.). S.D.F. was supported by a King's College junior research fellowship. We are grateful to Mark Skehel and the MRC-LMB Mass spectrometry service for protein characterization and the LFQ-based proteomics, and to Kate Heesom at Bristol University for TMT label-based proteomics.

Received: March 23, 2016

Revised: May 13, 2016

Accepted: May 23, 2016

Published: July 21, 2016

REFERENCES

- Agard, N.J., Prescher, J.A., and Bertozzi, C.R. (2004). A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **126**, 15046–15047.
- Agarwal, P., Beahm, B.J., Shieh, P., and Bertozzi, C.R. (2015). Systemic fluorescence imaging of zebrafish glycans with bioorthogonal chemistry. *Angew. Chem. Int. Ed Engl.* **54**, 11504–11510.
- Asare-Okai, P.N., Agustin, E., Fabris, D., and Royzen, M. (2014). Site-specific fluorescence labelling of RNA using bio-orthogonal reaction of trans-cyclooctene and tetrazine. *Chem. Commun. (Camb.)* **50**, 7844–7847.
- Bagert, J.D., van Kessel, J.C., Sweredoski, M.J., Feng, L., Hess, S., Bassler, B.L., and Tirrell, D.A. (2016). Time-resolved proteomic analysis of quorum sensing in *Chem. Sci.* **7**, 1797–1806.
- Bantscheff, M., Lemeer, S., Savitski, M.M., and Kuster, B. (2012). Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal. Bioanal. Chem.* **404**, 939–965.
- Blackman, M.L., Royzen, M., and Fox, J.M. (2008). Tetrazine ligation: fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. *J. Am. Chem. Soc.* **130**, 13518–13519.
- Bormann, A., Milles, S., Plass, T., Dommerholt, J., Verkade, J.M., Wiessler, M., Schultz, C., van Hest, J.C., van Delft, F.L., and Lemke, E.A. (2012). Genetic encoding of a bicyclo[6.1.0]nonyne-charged amino acid enables fast cellular protein imaging by metal-free ligation. *Chembiochem* **13**, 2094–2099.
- Busskamp, H., Batroff, E., Niederwieser, A., Abdel-Rahman, O.S., Winter, R.F., Wittmann, V., and Marx, A. (2014). Efficient labelling of enzymatically synthesized vinyl-modified DNA by an inverse-electron-demand Diels-Alder reaction. *Chem. Commun.* **50**, 10827–10829.
- Chesarino, N.M., Hach, J.C., Chen, J.L., Zaro, B.W., Rajaram, M.V., Turner, J., Schlesinger, L.S., Pratt, M.R., Hang, H.C., and Yount, J.S. (2014). Chemoproteomics reveals Toll-like receptor fatty acylation. *BMC Biol.* **12**, 91.
- Chin, J.W. (2014). Expanding and reprogramming the genetic code of cells and animals. *Annu. Rev. Biochem.* **83**, 379–408.
- Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell Proteomics* **13**, 2513–2526.
- Cravatt, B.F., Wright, A.T., and Kozarich, J.W. (2008). Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu. Rev. Biochem.* **77**, 383–414.
- Denk, C., Svatoněk, D., Filip, T., Wanek, T., Lumpi, D., Frohlich, J., Kuntner, C., and Mikula, H. (2014). Development of a (18)F-labeled tetrazine with favorable pharmacokinetics for bioorthogonal PET imaging. *Angew. Chem. Int. Ed Engl.* **53**, 9655–9659.
- Devaraj, N.K., Weissleder, R., and Hilderbrand, S.A. (2008). Tetrazine-based cycloadditions: application to pretargeted live cell imaging. *Bioconjug. Chem.* **19**, 2297–2299.
- Dieterich, D.C., Link, A.J., Graumann, J., Tirrell, D.A., and Schuman, E.M. (2006). Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc. Natl. Acad. Sci. USA* **103**, 9482–9487.
- Dieterich, D.C., Lee, J.J., Link, A.J., Graumann, J., Tirrell, D.A., and Schuman, E.M. (2007). Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging. *Nat. Protoc.* **2**, 532–540.
- Dommerholt, J., van Rooijen, O., Bormann, A., Guerra, C.F., Bickelhaupt, F.M., and van Delft, F.L. (2014). Highly accelerated inverse electron-demand cycloaddition of electron-deficient azides with aliphatic cyclooctynes. *Nat. Commun.* **5**, 5378.
- Elliott, T.S., Bianco, A., and Chin, J.W. (2014a). Genetic code expansion and bioorthogonal labelling enables cell specific proteomics in an animal. *Curr. Opin. Chem. Biol.* **27**, 154–160.
- Elliott, T.S., Townsley, F.M., Bianco, A., Ernst, R.J., Sachdeva, A., Elsasser, S.J., Davis, L., Lang, K., Pisa, R., Greiss, S., et al. (2014b). Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. *Nat. Biotechnol.* **32**, 465–472.
- Erdmann, I., Marter, K., Kobler, O., Niehues, S., Abele, J., Muller, A., Bussmann, J., Storkebaum, E., Ziv, T., Thomas, U., et al. (2015). Cell-selective labelling of proteomes in *Drosophila melanogaster*. *Nat. Commun.* **6**, 7521.
- Hong, V., Presolski, S.I., Ma, C., and Finn, M.G. (2009). Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew. Chem. Int. Ed Engl.* **48**, 9879–9883.
- Kaya, E., Vrabl, M., Deiml, C., Prill, S., Fluxa, V.S., and Carell, T. (2012). A genetically encoded norbornene amino acid for the mild and selective modification of proteins in a copper-free click reaction. *Angew. Chem. Int. Ed Engl.* **51**, 4466–4469.
- Klick, K.L., Saxon, E., Tirrell, D.A., and Bertozzi, C.R. (2002). Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. USA* **99**, 19–24.
- Kleiner, R.E., Verma, P., Molloy, K.R., Chait, B.T., and Kapoor, T.M. (2015). Chemical proteomics reveals a gammaH2AX-53BP1 interaction in the DNA damage response. *Nat. Chem. Biol.* **11**, 807–814.
- Kurra, Y., Odoi, K.A., Lee, Y.J., Yang, Y., Lu, T., Wheeler, S.E., Torres-Kolbus, J., Deiters, A., and Liu, W.R. (2014). Two rapid catalyst-free click reactions for

- in vivo protein labeling of genetically encoded strained alkene/alkyne functionalities. *Bioconjug. Chem.* 25, 1730–1738.
- Lang, K., and Chin, J.W. (2014a). Bioorthogonal reactions for labeling proteins. *ACS Chem. Biol.* 9, 16–20.
- Lang, K., and Chin, J.W. (2014b). Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem. Rev.* 114, 4764–4806.
- Lang, K., Davis, L., Torres-Kolbus, J., Chou, C., Deiters, A., and Chin, J.W. (2012a). Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat. Chem.* 4, 298–304.
- Lang, K., Davis, L., Wallace, S., Mahesh, M., Cox, D.J., Blackman, M.L., Fox, J.M., and Chin, J.W. (2012b). Genetic Encoding of bicyclononynes and trans-cyclooctenes for site-specific protein labeling in vitro and in live mammalian cells via rapid fluorogenic Diels-Alder reactions. *J. Am. Chem. Soc.* 134, 10317–10320.
- Liu, C.C., and Schultz, P.G. (2010). Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* 79, 413–444.
- Mahdavi, A., Hamblin, G.D., Jindal, G.A., Bagert, J.D., Dong, C., Sweredoski, M.J., Hess, S., Schuman, E.M., and Tirrell, D.A. (2016). Engineered aminoacyl-tRNA synthetase for cell-selective analysis of mammalian protein synthesis. *J. Am. Chem. Soc.* 138, 4278–4281.
- Mancuso, L., Jurjens, G., Hermans, J., Harmrolfs, K., Eichner, S., Fohrer, J., Collisi, W., Sasse, F., and Kirschning, A. (2013). Bioreduction of aryl azides during mutasynthesis of new ansamitocins. *Org. Lett.* 15, 4442–4445.
- McAlister, G.C., Huttlin, E.L., Haas, W., Ting, L., Jedrychowski, M.P., Rogers, J.C., Kuhn, K., Pike, I., Grothe, R.A., Blethrow, J.D., et al. (2012). Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal. Chem.* 84, 7469–7478.
- Ngo, J.T., and Tirrell, D.A. (2011). Noncanonical amino acids in the interrogation of cellular protein synthesis. *Acc. Chem. Res.* 44, 677–685.
- Ngo, J.T., Schuman, E.M., and Tirrell, D.A. (2013). Mutant methionyl-tRNA synthetase from bacteria enables site-selective N-terminal labeling of proteins expressed in mammalian cells. *Proc. Natl. Acad. Sci. USA* 110, 4992–4997.
- Nikic, I., Plass, T., Schraidt, O., Szymanski, J., Briggs, J.A., Schultz, C., and Lemke, E.A. (2014). Minimal tags for rapid dual-color live-cell labeling and super-resolution microscopy. *Angew. Chem. Int. Ed Engl.* 53, 2245–2249.
- Patterson, D.M., Nazarova, L.A., Xie, B., Kamber, D.N., and Prescher, J.A. (2012). Functionalized cyclopropanes as bioorthogonal chemical reporters. *J. Am. Chem. Soc.* 134, 18638–18643.
- Patterson, D.M., Jones, K.A., and Prescher, J.A. (2014). Improved cyclopropane reporters for probing protein glycosylation. *Mol. Biosyst.* 10, 1693–1697.
- Plass, T., Milles, S., Koehler, C., Szymanski, J., Mueller, R., Wiessler, M., Schultz, C., and Lemke, E.A. (2012). Amino acids for Diels-Alder reactions in living cells. *Angew. Chem. Int. Ed Engl.* 51, 4166–4170.
- Prescher, J.A., and Bertozzi, C.R. (2006). Chemical technologies for probing glycans. *Cell* 126, 851–854.
- Pyka, A.M., Domnick, C., Braun, F., and Kath-Schorr, S. (2014). Diels-Alder cycloadditions on synthetic RNA in mammalian cells. *Bioconjug. Chem.* 25, 1438–1443.
- Rieder, U., and Luedtke, N.W. (2014). Alkene-tetrazine ligation for imaging cellular DNA. *Angew. Chem. Int. Ed Engl.* 53, 9168–9172.
- Sachdeva, A., Wang, K., Elliott, T., and Chin, J.W. (2014). Concerted, rapid, quantitative, and site-specific dual labeling of proteins. *J. Am. Chem. Soc.* 136, 7785–7788.
- Selvaraj, R., Giglio, B., Liu, S., Wang, H., Wang, M., Yuan, H., Chintala, S.R., Yap, L.P., Conti, P.S., Fox, J.M., et al. (2015). Improved metabolic stability for 18F PET probes rapidly constructed via tetrazine trans-cyclooctene ligation. *Bioconjug. Chem.* 26, 435–442.
- Shih, H.W., Kamber, D.N., and Prescher, J.A. (2014). Building better bioorthogonal reactions. *Curr. Opin. Chem. Biol.* 21, 103–111.
- Smekens, J.M., Chen, W., and Wu, R. (2015). Mass spectrometric analysis of the cell surface N-glycoproteome by combining metabolic labeling and click chemistry. *J. Am. Soc. Mass Spectrom.* 26, 604–614.
- Szychowski, J., Mahdavi, A., Hodas, J.J., Bagert, J.D., Ngo, J.T., Landgraf, P., Dieterich, D.C., Schuman, E.M., and Tirrell, D.A. (2010). Cleavable biotin probes for labeling of biomolecules via azide-alkyne cycloaddition. *J. Am. Chem. Soc.* 132, 18351–18360.
- Thompson, A., Schafer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A.K., and Hamon, C. (2003). Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* 75, 1895–1904.
- Tsai, Y.H., Essig, S., James, J.R., Lang, K., and Chin, J.W. (2015). Selective, rapid and optically switchable regulation of protein function in live mammalian cells. *Nat. Chem.* 7, 554–561.
- Uttamapinant, C., Howe, J.D., Lang, K., Beranek, V., Davis, L., Mahesh, M., Barry, N.P., and Chin, J.W. (2015). Genetic code expansion enables live-cell and super-resolution imaging of site-specifically labeled cellular proteins. *J. Am. Chem. Soc.* 137, 4602–4605.
- Vanbeselaere, J., Chang, L.Y., Harduin-Lepers, A., Fabre, E., Yamakawa, N., Slomianny, C., Biot, C., Khoo, K.H., and Guerardel, Y. (2012). Mapping the expressed glycome and glycosyltransferases of zebrafish liver cells as a relevant model system for glycosylation studies. *J. Proteome Res.* 11, 2164–2177.
- Verhelst, S.H., Fonovic, M., and Bogoy, M. (2007). A mild chemically cleavable linker system for functional proteomic applications. *Angew. Chem. Int. Ed Engl.* 46, 1284–1286.
- Wang, Q., Chan, T.R., Hilgraf, R., Fokin, V.V., Sharpless, K.B., and Finn, M.G. (2003). Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition. *J. Am. Chem. Soc.* 125, 3192–3193.
- Wang, M., Weiss, M., Simonovic, M., Haertinger, G., Schrimpf, S.P., Hengartner, M.O., and von Mering, C. (2012). PaxDb, a database of protein abundance averages across all three domains of life. *Mol. Cell Proteomics* 11, 492–500.
- Wang, M., Herrmann, C.J., Simonovic, M., Szklarczyk, D., and von Mering, C. (2015). Version 4.0 of PaxDb: protein abundance data, integrated across model organisms, tissues, and cell-lines. *Proteomics* 15, 3163–3168.
- Werner, T., Becher, I., Sweetman, G., Doce, C., Savitski, M.M., and Bantscheff, M. (2012). High-resolution enabled TMT 8-plexing. *Anal. Chem.* 84, 7188–7194.
- Yang, Y.Y., Grammel, M., Raghavan, A.S., Charron, G., and Hang, H.C. (2010). Comparative analysis of cleavable azobenzene-based affinity tags for bioorthogonal chemical proteomics. *Chem. Biol.* 17, 1212–1222.
- Yang, J., Seckute, J., Cole, C.M., and Devaraj, N.K. (2012). Live-cell imaging of cyclopropane tags with fluorogenic tetrazine cycloadditions. *Angew. Chem. Int. Ed Engl.* 51, 7476–7479.
- Yuet, K.P., and Tirrell, D.A. (2014). Chemical tools for temporally and spatially resolved mass spectrometry-based proteomics. *Ann. Biomed. Eng.* 42, 299–311.
- Yuet, K.P., Doma, M.K., Ngo, J.T., Sweredoski, M.J., Graham, R.L., Moradian, A., Hess, S., Schuman, E.M., Sternberg, P.W., and Tirrell, D.A. (2015). Cell-specific proteomic analysis in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* 112, 2705–2710.
- Zhang, M.M., Wu, P.Y., Kelly, F.D., Nurse, P., and Hang, H.C. (2013). Quantitative control of protein S-palmitoylation regulates meiotic entry in fission yeast. *PLoS Biol.* 11, e1001597.
- Zheng, T., Jiang, H., and Wu, P. (2013). Single-stranded DNA as a cleavable linker for bioorthogonal click chemistry-based proteomics. *Bioconjug. Chem.* 24, 859–864.